Are male births more likely than female births? The first p-value.

Unit 3 Lecture 1

Jonathan Auerbach
STAT 490 Pre-Cap Prof Development
jauerba@gmu.edu

October 14, 2021

How do scientists determine whether data support a theory?

These slides use the following R packages

Setup:

```
library("knitr")
library("HistData")
library("tidyverse")
theme_set(theme_bw())
```

Are male births more likely than female births?

- Arbuthnot (1710) retrieved 82 years of London christenings (1629-1710)
- The number of boys exceeded the number of girls every year
- Arbuthnot reasoned: were birth rates equal, the probability of more boys each year = probability a fair coin lands on heads 82 times in a row
- This probability (the p-value) is essentially zero

John Arbuthnot (1722)

An argument for divine providence (1710)

II. *An Argument for Divine Providence, taken from the conſtant Regularity obſerv'd in the Births of both Sexes. By Dr.* John Arbuthnott, *Phyſitian in Ordinary to Her Majeſty, and Fellow of the College of Phyſitians and the Royal Society.*

AMong innumerable Footſteps of Divine Providence to be found in the Works of Nature, there is a very remarkable one to be obſerved in the exact Ballance that is maintained, between the Numbers of Men and Women; for by this means it is provided, that the Species may never fail, nor periſh, ſince every Male may have its Female, and of a proportionable Age. This Equality of Males and Females is not the Effect of Chance but Divine Providence, working for a good End, which I thus demonſtrate:

Let there be a Die of Two ſides, M and F, (which denote Croſs and Pile), now to find all the Chances of any determinate Number of ſuch Dice, let the Binome M+F be raiſed to the Power, whoſe Exponent is the Number of Dice given; the Coefficients of the Terms will ſhew all the Chances ſought. For Example, in Two Dice of Two ſides M+F the Chances are $M^2+2\,MF+F^2$, that is, One Chance for M double, One for F double, and Two for M ſingle and F ſingle; in Four ſuch Dice there are Chances $M^4+4\,M^3F+6\,M^2F^2+4\,MF^3+F^4$, that is, One Chance for M quadruple, One for F quadruple, Four for triple M and ſingle F, Four for ſingle M and triple F, and Six for M double and F double; and univerſally, if the Number of Dice be n, all their Chances will be expreſſed in this Series

M^n+

leſs than any aſſignable Fraction. From whence it follows, that it is Art, not Chance, that governs.

There ſeems no more probable Cauſe to be aſſigned in Phyſicks for this Equality of the Births, than that in our firſt Parents Seed there were at firſt formed an equal Number of both Sexes.

Scholium. From hence it follows, that Polygamy is contrary to the Law of Nature and Juſtice, and to the Propagation of Human Race; for where Males and and Females are in equal number, if one Man takes Twenty Wives, Nineteen Men muſt live in Celibacy, which is repugnant to the Deſign of Nature; nor is it probable that Twenty Women will be ſo well impregnated by one Man as by Twenty.

Chriſtened.			Chriſtened.		
Anno.	*Males.*	*Females.*	*Anno.*	*Males.*	*Females.*
1629	5218	4683	1648	3363	3181
30	4858	4457	49	3079	2746
31	4422	4102	50	2890	2722
32	4994	4590	51	3231	2840
33	5158	4839	52	3220	2908
34	5035	4820	53	3196	2959
35	5106	4928	54	3441	3179
36	4917	4605	55	3655	3349
37	4703	4457	56	3668	3382
38	5359	4952	57	3396	3289
39	5366	4784	58	3157	3013
40	5518	5332	59	3209	2781
41	5470	5200	60	3724	3247
42	5460	4910	61	4748	4107
43	4793	4617	62	5216	4803
44	4107	3997	63	5411	4881
45	4047	3919	64	6041	5681
46	3768	3395	65	5114	4858
47	3796	3536	66	4678	4319

Arbuthnot's Data

```
Arbuthnot %>%
  select(Year, Males, Females) %>%
  filter(Year < 1634 | Year > 1707) %>%
  kable()
```

Year	Males	Females
1629	5218	4683
1630	4858	4457
1631	4422	4102
1632	4994	4590
1633	5158	4839
1708	8239	7623
1709	7840	7380
1710	7640	7288

Sign Test

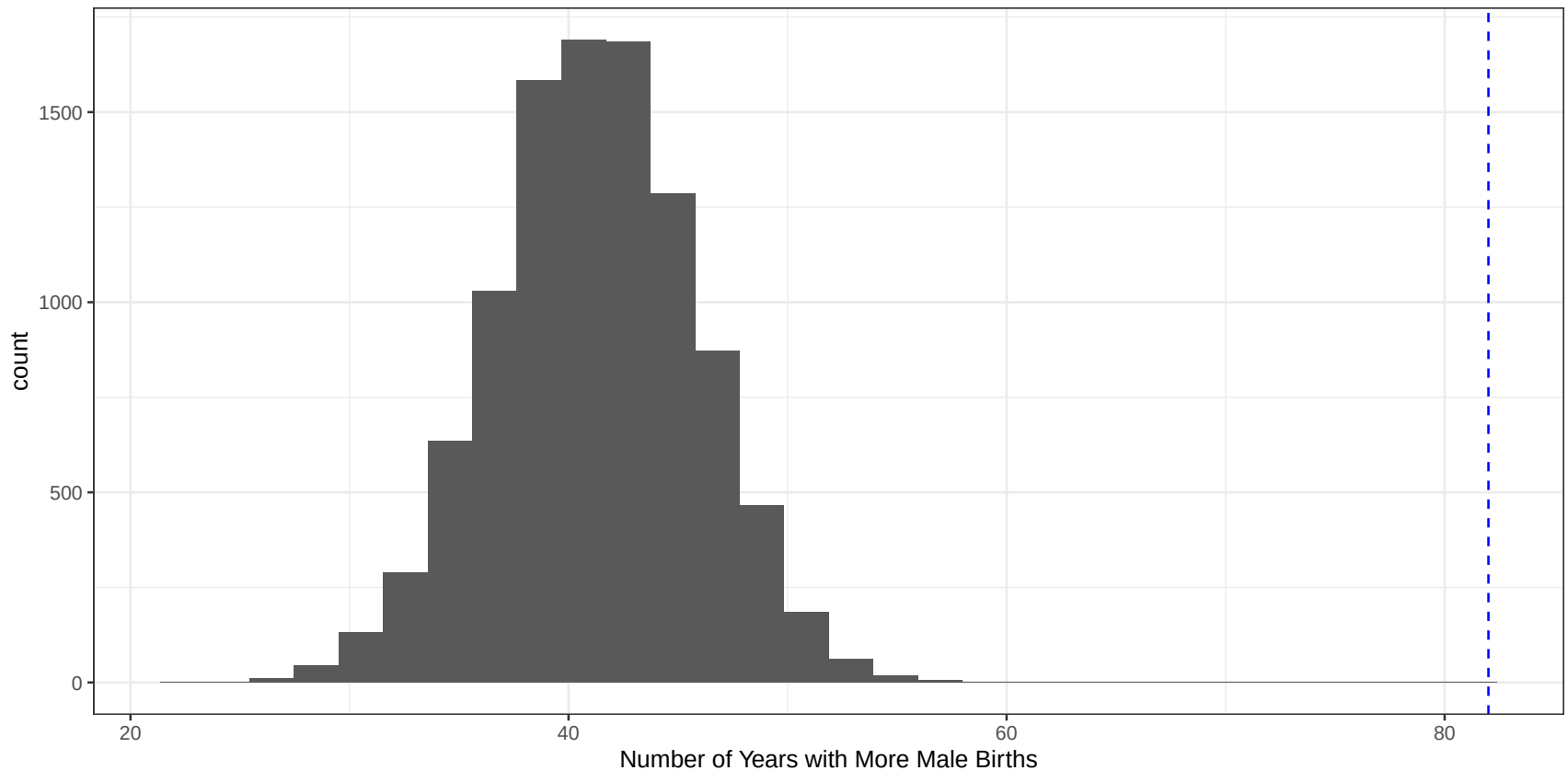
```
Arbuthnot %>%
 mutate(Heads = ifelse(Males - Females > 0, 1, 0)) %>%
 summarize(Num_Heads  = sum(Heads),
           Num_Trials = n()) %>%
 transmute(
   `p value` =
     binom.test(x = Num_Heads,
                n = Num_Trials,
                p = .5,
                alternative = "greater")$p.value) %>%
 kable(digits = Inf)
```

p value
2.067952e-25

This is the same as $\frac{1}{2^{82}} = 2.0679515 \times 10^{-25}$.

Simulations of Sign Test under Null Hypothesis

```
ggplot(tibble(sims = rbinom(1e4, 82, .5))) +
  geom_histogram(aes(x = sims)) +
  geom_vline(xintercept = 82,
             color = "blue", linetype = 2) +
  labs(x = "Number of Years with More Male Births")
```



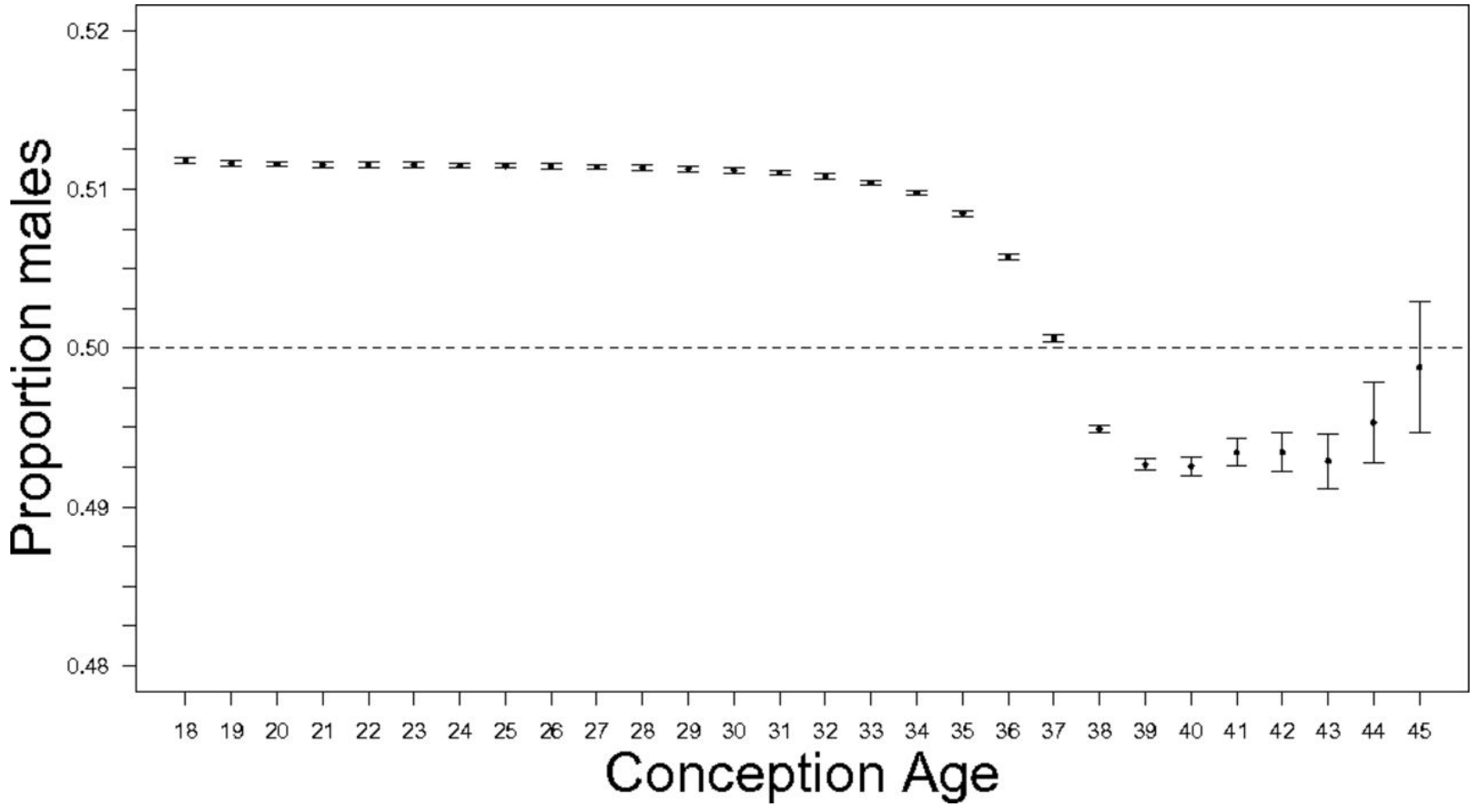
Why are male births more likely than female births?

- Arbuthnot thought the difference was due to a wise creator carefully adjusting for the risk men face hunting.
- This theory is not supported by the data. Arbuthnot proved the difference in the birth rates is not zero. He did not prove it is consistent with the risk men face hunting.
- To support his theory, Arbuthnot might have determined the risk men face hunting, and then tested whether this rate is consistent with the excess of male over female births.

Why are male births more likely than female births?

- Recent research suggests the birth ratio is balanced at conception (Figure from Orzack et al (2015))

- Female embryos more likely to be lost during pregnancy



References

1. Arbuthnot, John. "An Argument for Divine Providence." Philosophical Transactions 27 (1710): 186-190.

2. Friendly, Michael. "HistData: Data sets from the history of statistics and data visualization." R package version 0.7-5 (2014).

3. Orzack, Steven Hecht, et al. "The human sex ratio from conception to birth." Proceedings of the National Academy of Sciences (2015): 201416546.

4. Stigler, Stephen. "The seven pillars of statistical wisdom." Harvard University Press, 2016.